Սարգիս Անդրեասի Հովհաննիսյան

ՕԲՅԵԿՏՆԵՐԻ ՀԱՅՏՆԱԲԵՐՈՒՄ ԱՆԲԱՐԵՆՊԱՍՏ ԵՂԱՆԱԿԱՅԻՆ ՊԱՅՄԱՆՆԵՐՈՒՄ՝ ՕԳՏԱԳՈՐԾԵԼՈՎ ԽՈՐԸ ՈՒՍՈՒՑՈՒՄ ԵՎ ՋԵՐՄԱՅԻՆ-ՏԵՍԱՆԵԼԻ ՊԱՏԿԵՐՆԵՐ

Ե.13.05 - «Մաթեմատիկական մոդելավորում, թվային մեթոդներ և ծրագրերի համալիրներ» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի համար

ՍԵՂՄԱԳԻՐ

Երևան 2025

Sargis Hovhannisyan

OBJECT DETECTION IN ADVERSE WEATHER USING DEEP LEARNING AND THERMAL-VISIBLE IMAGING

SYNOPSYS

of the dissertation for obtaining a Ph.D. degree in Technical Sciences on specialty 05.13.05 "Mathematical modelling, numerical methods and program complexes"

Yerevan 2025

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտում:

Գիտական ղեկավար`	ֆիզ. մաթ. գիտ. դոկտոր **Ս. Ս. Աղայան**
Պաշտոնական ընդդիմախոսներ`	ֆիզ. մաթ. գիտ. դոկտոր **Լ. Հ. Ասլանյան**
	ֆիզ. մաթ. գիտ. թեկնածու **Կ. Ի. Կարապետյան**
Առաջատար կազմակերպություն`	Հայաստանի ազգային պոլիտեխնիկական համալսարան

Ատենախոսության պաշտպանությունը տեղի կունենա 2025թ. հուլիսի 10-ին ժամը 13:00-ին` ՀՀ ԳԱԱ Ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտի մասնագիտացված խորհուրդ 037 «Ինֆորմատիկա» մասնագիտական խորհրդի նիստում հետևյալ հասցեով` Երևան, 0014, Պ. Սևակ փող. 1:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ ԻԱՊԻ գրադարանում:

Սեղմագիրն առաքված է 2025թ . հունիսի 10-ին:

Մասնագիտական խորհրդի գիտական քարտուղար ֆիզ. մաթ. գիտ. դոկտոր`   **Մ. Ե. Հարությունյան**

The topic of the dissertation was approved at the Institute of Informatics and Automation Problems of NAS RA.

Scientific supervisor:	Doctor of phys-math sciences **S. S. Agaian**
Official opponents:	Doctor of phys-math sciences **L. H. Aslanyan**
	Candidate of phys-math sciences **K. I. Karapetyan**
Leading organization:	National Polytechnic University of Armenia

The dissertation defense will take place on July 10, 2025, at 13:00, at the Specialized Council 037 «Informatics» at the Institute of Informatics and Automation Problems of NAS RA. Address: 1 P. Sevak Street, Yerevan 0014.

The dissertation is available at the library of ПAP NAS RA.

The abstract is delivered on 10 June, 2025.

Scientific Secretary of the Specialized Council Doctor of phys-math sciences   **M. E. Harutyunyan**

Relevance of the Research

Object detection is identifying and localizing objects within images or video frames, typically marked by bounding boxes indicating the location and class of detected objects (Figure 1 illustrates an example of object detection output). Accurate object detection is critical across many practical domains, impacting safety, efficiency, and security. In Autonomous Vehicles, object detection supports functionalities including collision avoidance, pedestrian detection, lane identification, and traffic sign recognition. In Unmanned Aerial Vehicles (UAV) systems, object detection enables route planning, obstacle avoidance, and target monitoring. Accurate detection and identification of wildlife species in automated monitoring systems facilitate improved ecological monitoring, conservation efforts, and anti-poaching initiatives. In smart-city surveillance, object detection improves urban safety through better monitoring of traffic incidents, crime detection, and public safety management. In industrial inspection systems, object detection supports critical tasks such as defect detection, quality assurance, and safety inspections in industrial environments.
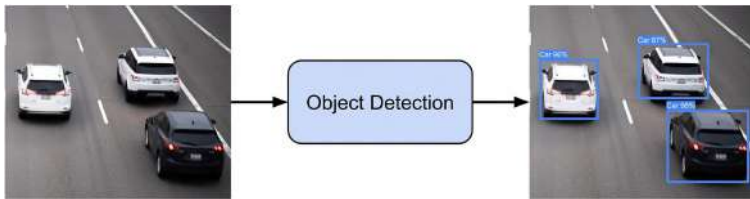

Figure 1: Example of an object detection output.

These systems utilized two types of imaging technologies: Red Green Blue (RGB) visible and Thermal Infrared (TIR) cameras. RGB cameras capture visible light across the red, green, and blue spectrum, providing high-resolution color imagery with excellent detail in well-lit conditions. TIR cameras detect heat signatures by capturing emitted infrared radiation from objects, enabling object detection regardless of lighting conditions and allowing temperature-based differentiation.

While current detection systems achieve human-level precision on benchmark datasets and are widely deployed in commercial products due to deep-learning models that learn visual patterns directly from data, they are typically optimized for clear-weather images with good lighting conditions. Despite these strengths, these systems are significantly degraded under adverse weather conditions. Numerous challenges persist that limit the reliability and effectiveness of these systems in real-world applications. Studies have shown that state-of-the-art (SOTA) detectors' accuracy can drop by around **30-40%** as haze density increases.

RGB sensors are inherently dependent on ambient lighting and highly susceptible to visual degradation caused by fog, haze, shadows, and nighttime darkness. Although TIR sensors function effectively in darkness, through light fog, and under headlight glare, they suffer from inherent limitations, including low contrast, weak edge definition, and sensor noise that complicate accurate object detection.

Even thermal imaging, while robust in low-visibility scenarios, can be compromised by severe atmospheric interference such as heavy fog or haze, which reduces contrast and edge sharpness, further impairing detection accuracy. Thermal video introduces additional complications, including motion blur and camera jitter, making object detection even more challenging in dynamic environments where
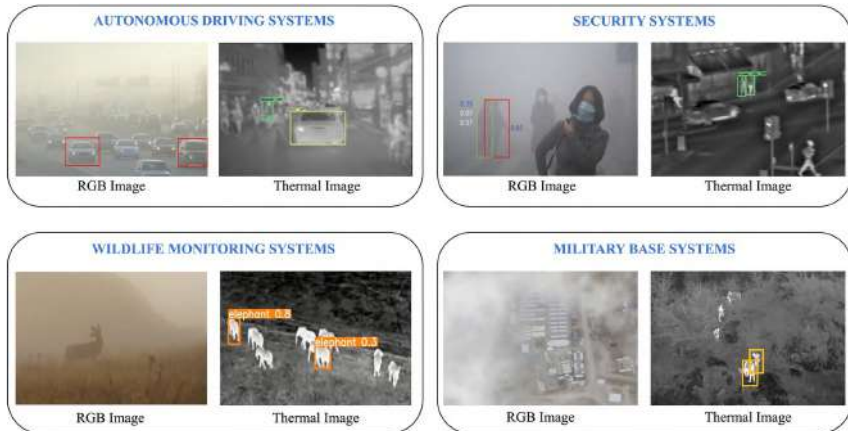
3

Figure 2: Examples of object detection failures under bad weather conditions across different imaging modalities.

temporal consistency becomes critical for reliable performance (see Figure 2 for examples of such failures).

These detection failures translate directly into real-world harms with significant consequences. When object detection systems fail due to environmental degradations such as fog, haze, poor illumination, or occlusions, critical image features become obscured, resulting in missed detections or misclassifications that can have severe implications.

- Road safety reports by the United States (U.S.) indicate that low-visibility conditions, including fog, haze, and nighttime driving, contribute to a disproportionate number of fatal crashes. Although these conditions account for a small percentage of total driving exposure, they collectively account for nearly **50%** of all traffic fatalities in the U.S.. The increased risk results from reduced driver awareness, delayed reaction times, and reduced effectiveness of vehicle safety systems. Studies further illustrate that the effectiveness of autonomous braking systems can be reduced by 30% to 80% when visibility is reduced due to severe fog conditions, significantly narrowing reaction times and increasing collision risks.

- Surveillance systems at fixed locations also experience severe impairments under low visibility conditions. For instance, camera-based motion detection systems frequently generate false alarms triggered by fog, dust, and insects. According to U.S. policing studies, false alarm rates for burglar alarm dispatches range from **94% to 98%**, placing unnecessary burdens on law enforcement resources. Furthermore, criminals exploit visibility impairments, with analyses indicating that approximately **50%** of residential burglaries occur during nighttime or under low visibility when RGB camera systems demonstrate the lowest reliability.

- Wildlife conservation efforts are similarly impacted, as around **80%** of unauthorized wildlife hunting incidents occur under dense atmospheric haze. Although drone surveillance provides crucial monitoring capabilities, aerial operations frequently encounter significant disruptions

4

from fog and dust, limiting operational hours substantially. For example, a U.S. border-security audit showed that visibility restrictions kept unmanned aircraft airborne for about **22%** of their scheduled hours.

- Aviation safety and military operations routinely face "degraded visual environments", including conditions induced by fog, dust, and smoke. According to U.S. Army safety analyses, disorientation or obstacle collisions account for approximately **24%** of helicopter crashes and **44%** of fatalities. Additionally, commercial airports face operational disruptions due to fog, with major airports operating under instrument flight rules between **15% to 23%** of annual operational hours, reinforcing the critical need for effective visual enhancement solutions.
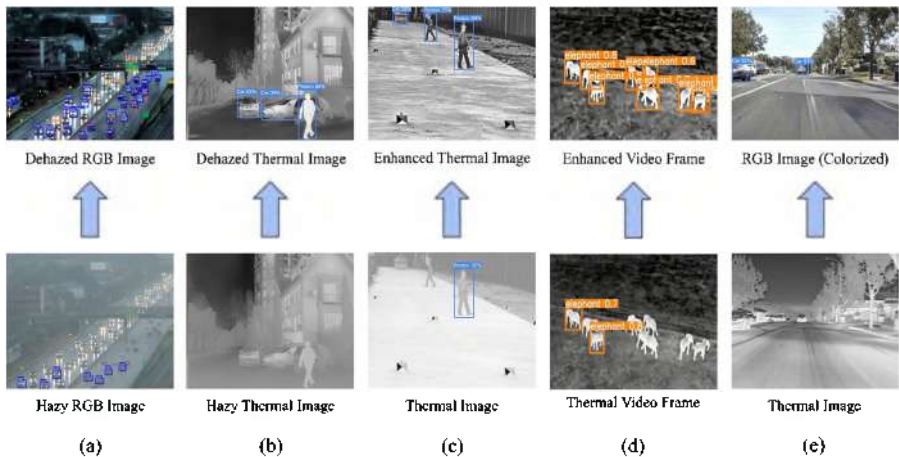

Dehazed RGB Image   Dehazed Thermal Image   Enhanced Thermal Image   Enhanced Video Frame   RGB Image (Colorized)

Hazy RGB Image   Hazy Thermal Image   Thermal Image   Thermal Video Frame   Thermal Image

(a)   (b)   (c)   (d)   (e)

Figure 3: Impact of Enhancement Methods on Object Detection for RGB and TIR Imagery.

Several promising methodological approaches have emerged to address the challenges of reliable object detection under adverse conditions. Image dehazing represents a foundational approach to improving visibility in degraded visual conditions by removing atmospheric interference such as haze or fog from images (see Figure 3 (a)). This technique addresses the significant impairment of computer vision algorithms caused by reduced visibility, contrast, and detail clarity in adverse weather. Recent deep learning approaches show significant advancements by automatically learning important features from data; however, they face generalization limitations under varying real-world conditions and struggle with complex backgrounds or atmospheric degradation.

Similarly, TIR dehazing addresses the unique degradation patterns affecting thermal imagery in adverse weather conditions. While thermal cameras offer inherent advantages in low-light conditions, they remain susceptible to quality reduction from heavy atmospheric interference. Thermal images experience distinct degradation characteristics, including diminished contrast, edge blurring, and decreased clarity resulting from atmospheric scattering and absorption (see Figure 3 (b)). Contemporary approaches leverage Convolutional Neural Networks (CNN) to enhance thermal image quality. However, significant challenges persist regarding limited dataset availability, the fidelity of synthetic data

5

generation, and model adaptability across diverse thermal imaging systems with varying specifications.

General thermal image enhancement techniques aim to improve the overall quality and interpretability of thermal imagery by addressing common challenges, including low contrast, detail obscuration, ghosting effects from overlapping thermal radiation, and inconsistent sensor characteristics (see Figure 3 (c)). Recent approaches, including GAN-based approaches and CNN architectures, have advanced image quality considerably but continue to face generalization difficulties in complex scenarios, particularly those involving reflective materials and ambiguous thermal patterns.

Thermal video enhancement extends beyond static image processing by addressing the temporal characteristics inherent to thermal video data. This introduces additional complexities, such as motion blur, temporal inconsistencies, rapid scene dynamics, and variability in sensor responses over time (see Figure 3 (d)). Methods that leverage temporal context encounter significant challenges in effectively handling non-rigid motion and complex thermal variations across frames. These temporal artifacts and inconsistencies severely undermine object detection performance, leading to unreliable detections, increased false positives, and missed targets, particularly in dynamic environments where consistent and accurate detection is essential.

Thermal image colorization represents a transformative approach to bridging the gap between TIR and RGB domains by addressing the inherent lack of color information and typically low contrast with unclear object boundaries in thermal imagery. TIR to RGB colorization is the process of transforming a single-channel TIR image into a three-channel color image that corresponds to visible-spectrum images. This process aims to generate realistic textures, colors, and visual details that would be present if the scene were captured by an RGB camera under favorable lighting conditions, as illustrated in Figure 3 (e). Colorizing thermal images enhances both human interpretability and compatibility with RGB-trained models. It provides a more intuitive visual representation and allows existing RGB-based algorithms to be applied to thermal data without extensive retraining. Recent deep learning frameworks for TIR-to-RGB translation encounter significant challenges, including semantic distortions, inconsistent detail preservation, temporal instability in video sequences, and suboptimal performance with small objects, which complicates accurate thermal-to-visible translation for precision applications like autonomous driving systems.

Given these substantial and persistent challenges, there remains a critical need for innovative approaches to enhance object detection systems' reliability under adverse weather and low-visibility conditions. This thesis addresses these challenges by proposing novel deep-learning methodologies tailored to improve image and video quality in visually degraded environments. In particular, we concentrate on advanced image dehazing methods that effectively mitigate atmospheric interference in RGB and TIR modalities. Furthermore, we investigate specialized enhancement techniques for thermal images and videos, employing recent advancements in neural architectures to handle unique degradations such as low contrast, edge ambiguity, and temporal inconsistencies. Finally, this work explores TIR-to-RGB colorization methods, bridging the gap between these imaging modalities to leverage RGB-based algorithms without extensive retraining, thus significantly improving object detection accuracy and reliability across practical, real-world scenarios.

Challenges of Object Detection

Reliable object detection remains challenging due to the inherent limitations and distinct vulnerabilities associated with different imaging modalities when operating in adverse environmental

6

conditions. While detection technologies are increasingly accurate in controlled or optimal conditions, their performance rapidly deteriorates when facing real-world scenarios involving degraded visual environments.

Despite **RGB imaging** widespread use, RGB-based detection methods inherently depend on ambient illumination and visibility conditions. Adverse scenarios such as fog, haze, heavy shadows, nighttime darkness, or noise introduced by bad weather significantly degrade RGB image quality, resulting in reduced contrast and loss of fine details critical for accurate detection. Consequently, detection reliability is severely affected, leading to frequent object mislocalization and misclassification. Furthermore, as most deep-learning detectors are trained predominantly on clear-weather images, their performance in challenging conditions is often compromised, highlighting the need for alternative imaging modalities that are more resilient to environmental impairments.

TIR imaging offers advantages for detection tasks by capturing emitted radiation rather than reflected light, providing resilience in varied lighting conditions. However, adoption is limited by several key challenges: *domain shift from RGB* data requiring specialized model training, as visible-spectrum trained models struggle with the fundamentally different visual features and contrast patterns in thermal data. *Low contrast and blurred edges* significantly reduce the detail visibility needed for accurate object detection, while sensor and spectral variability across different thermal imaging systems complicates model generalization across deployments. *Atmospheric degradation* from fog or haze affects image quality despite TIR's relative robustness, reducing contrast and blurring edges critical for detection accuracy. Additionally, *reflection artifacts* from surfaces like metal and glass create misleading contours that confuse detection algorithms. In video applications, *motion blur, camera instability, and temporal noise* accumulation further compromise detection consistency. These combined challenges significantly affect detection reliability in critical applications, requiring advanced enhancement techniques to ensure consistent performance across challenging operational environments.

The Goal of the Thesis

The goal of this thesis is to address the significant challenges posed by adverse weather conditions on object detection systems, developing accurate deep learning methodologies tailored to overcome these issues. The proposed approaches aim to surpass existing (SOTA) methods, achieving superior performance across multiple benchmark datasets and real-world scenarios. To accomplish these objectives, the thesis focuses on the following technical tasks:

1. Develop **dehazing frameworks** explicitly designed for **RGB** and **TIR** images.

2. Develop **thermal image and video enhancement** networks.

3. Develop a TIR-to-RGB **colorization pipeline** capable of translating thermal images into visually intuitive RGB representations.

4. Conduct meticulous evaluations of the proposed frameworks, assessing their performance using key metrics such as **detection accuracy** and **generalization capability**.
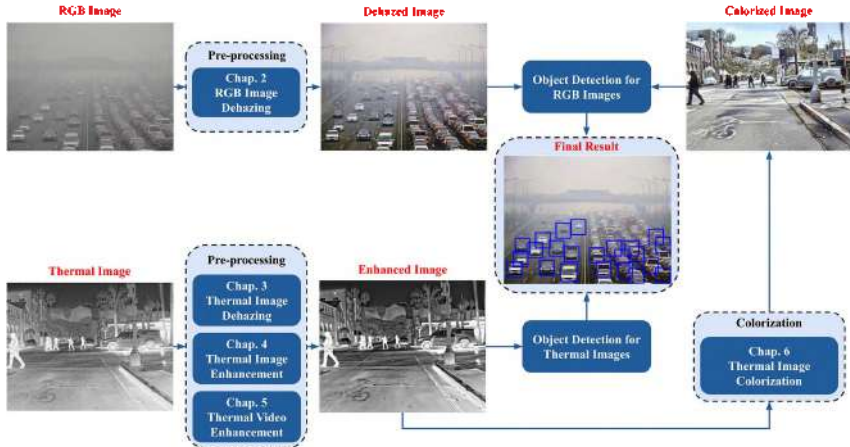
Figure 4: Overall Workflow of the Thesis Framework.

Structure of the Thesis

The dissertation consists of 7 chapters and a list of used literature. The thesis is written in 150 pages and has 249 literature references. The thesis contains 46 figures and 24 tables. Figure 4 illustrates the logical progression of this thesis, demonstrating how each chapter systematically contributes to a comprehensive framework for robust object detection under challenging visibility conditions.

Chapter 1 introduces the research context, clearly outlining existing limitations in object detection under adverse weather conditions. It articulates the research questions and objectives and highlights the key novel contributions made throughout the thesis.
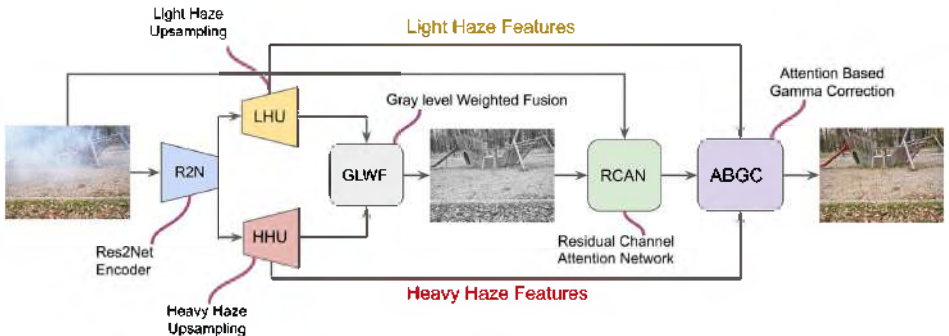

Figure 5: Overall architecture of EOD-Net.

Chapter 2 introduces an innovative deep-learning solution aimed at enhancing object detection in challenging weather conditions, particularly addressing significant degradation caused by haze and fog [1]. Despite considerable success, current state-of-the-art (SOTA) methods struggle with non-homogeneous haze, preserving natural colors, small training datasets, and generally poor adaptability

for downstream tasks such as object detection. To overcome these limitations, this chapter presents *EOD-Net* [2], a novel end-to-end RGB image dehazing architecture optimized specifically to improve object detection in hazy environments. The pipeline for EOD-Net, illustrated in Figure 5, features several key innovations: Firstly, images undergo dual-branch processing: one branch dedicated to light haze removal and another for heavy haze removal. These branches are merged via a specialized Gray-Level Weighted Fusion module utilizing the Dark Channel Prior, described for image J as:

$$J^{\text{dark}}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in r,g,b} J^c(y) \right) \tag{1}$$

Figure 6: Qualitative comparison on a real-world hazy image (top) and an outdoor test image (bottom).
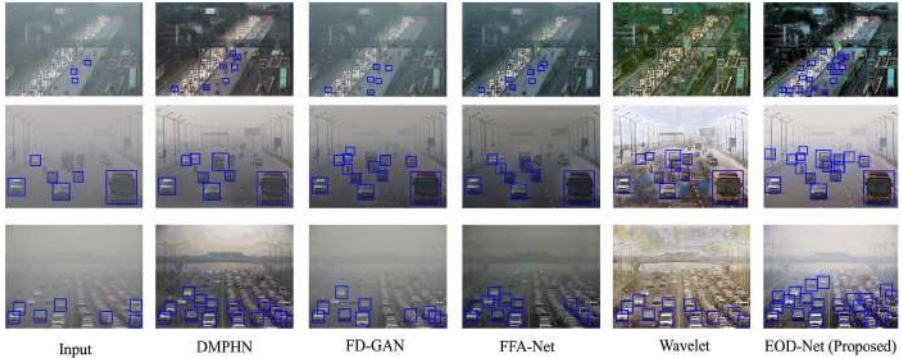

Figure 7: Vehicle detection results compared with SOTA dehazing methods.

This enhanced map enables the fusion of the dual branches into a unified, gray-level dehazed image. Subsequently, a specialized Residual Channel Attention Network handles the coloring step, restoring the natural color dynamics. Finally, an Attention-Based Gamma Correction (ABGC) module corrects any misleading colors and edges by applying a pixel-wise transformation defined as:

$$X^C_{\text{enhanced}} = \alpha(X^C)^\gamma + \beta \tag{2}$$

Here, α, β, and γ are outputs from the ABGC module, $C \in R, G, B$ denotes the color channel,

and X^C represents the input color channel. Comprehensive evaluations on synthetic datasets (I-Haze, O-Haze, NH-Haze2) and real-world datasets confirm EOD-Net's superior performance over existing methods across various image quality metrics. Figure 6 demonstrates qualitative comparisons on both hazy real-world and synthetic test set images, highlighting EOD-Net's capability to recover low-level and high-level details effectively in challenging scenarios. EOD-Net was evaluated against the top-performing methods (ADN, TBD, SRKTDN) from the NTIRE 2021 non-homogeneous image dehazing challenge, demonstrating superior performance in PSNR metrics and competitive results in SSIM on the NH-Haze2 dataset, as shown in Table 1. Furthermore, practical tests conducted on traffic surveillance footage reveal a substantial enhancement in object detection performance, shown in Figure 7. Unprocessed hazy footage typically detects only **3-4%** of vehicles, whereas EOD-Net improves detection to approximately **40%** in severely hazed conditions. These results underscore EOD-Net's potential to significantly enhance visual clarity and thus safety in critical real-world applications.

Table 1: Performance comparison on NH-Haze2 dataset

	ADN	TBD	SRKTDN	EOD-Net
PSNR	20.898	21.66	20.13	**23.953**
SSIM	**0.844**	0.843	0.803	0.842

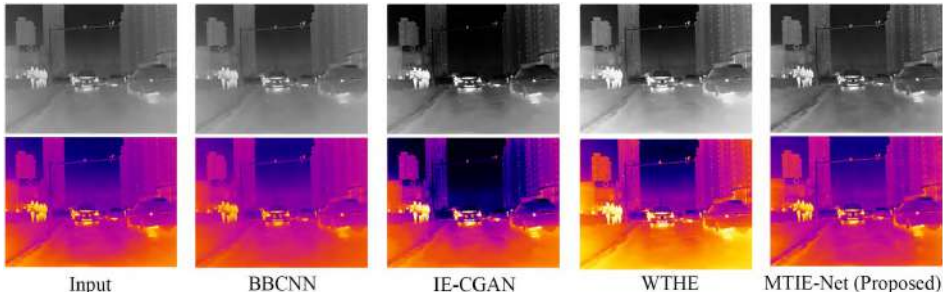| Input | BBCNN | IE-CGAN | WTHE | MTIE-Net (Proposed) |

Figure 8: Qualitative comparison of thermal image dehazing results and corresponding pseudo-color visualizations.

Chapter 3 addresses thermal image dehazing under severe atmospheric degradation caused by haze, smoke, and fog, which obscure details, lower contrast, and degrade downstream performance. Although current SOTA methods achieve acceptable results in thermal imaging under visible-light conditions, they still face considerable challenges in severe adverse conditions. To address these limitations, we propose *MTIE-Net* [3], a novel Mamba-based thermal image dehazing framework leveraging the Enhancement and Denoising State Space Model. This model integrates convolution-based attention mechanisms with state-space modeling, enabling simultaneous denoising and enhancement of thermal images, thereby restoring visibility while preserving crucial edges necessary for reliable object detection. Due to the lack of labeled hazy thermal images, we introduce a novel synthetic data generation pipeline leveraging the Atmospheric Scattering Model, defined as:

$$I(x) = J(x) \cdot t(x) + A \cdot (1 - t(x)), \quad \text{where} \quad t(x) = e^{-\beta d(x)} \tag{3}$$

with $J(x)$ as the haze-free image, A the atmospheric light, β the scattering coefficient, and $d(x)$ the depth map. Extensive evaluations conducted on the M3DF dataset demonstrate that MTIE-Net significantly outperforms traditional and deep-learning baselines across metrics including PSNR, SSIM, EME, BDIM, and MDIMTE. Qualitative comparisons in Figure 8 illustrate superior haze removal capabilities, avoiding the exposure artifacts of competing methods. Quantitative analysis in Table 2 further highlights the enhanced object detection performance on the M3FD dataset, with MTIE-Net achieving over a 35% improvement in detection accuracy compared to artificially generated hazy infrared images and approximately an 11% improvement over the best competing enhancement method under challenging conditions. MTIE-Net's robust performance and strong generalization to real-world scenarios make it particularly suitable for critical applications such as surveillance and safety-critical systems.

Table 2: Object detection evaluation ($mAP_{0.5}$) under adverse weather conditions.

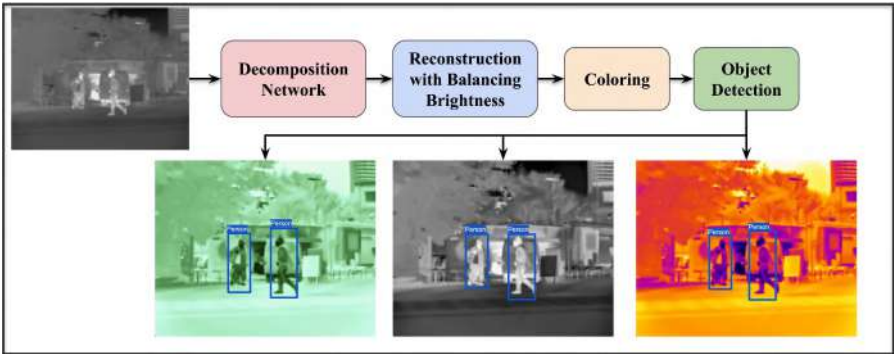| Measure	Day	Overcast	Night	Challenge	$mAP_{0.5}$
Hazy Infrared	0.718	0.721	0.620	0.618	0.710
Original Infrared	0.806	0.798	0.712	0.739	0.786
Visible	0.827	0.789	0.764	0.759	0.758
AGCCPF	0.811	0.799	0.739	0.748	0.789
BBCNN	0.815	0.805	0.743	0.747	0.790
IE-CGAN	0.816	0.808	0.768	0.754	0.791
WTHE	0.818	0.810	0.785	0.783	0.792
MTIE-Net	**0.828**	**0.819**	**0.849**	**0.871**	**0.812**

Figure 9: Overall pipeline of the proposed prior method.

Chapter 4 aims to develop an innovative physics-guided framework for thermal image enhancement, *PB-IID-Net*, addressing inherent challenges such as low contrast, ghosting artifacts, blurred edges, and sensor noise, which impair downstream vision tasks. Existing SOTA techniques often amplify noise and leave visible artifacts, limiting their practical applicability across different thermal imaging domains. This chapter enhances and extends the thermal image enhancement framework originally proposed in our prior work [4], illustrated in Figure 9. While the original concept laid the foundation for decomposition, this chapter further extends and refines the approach by incorporating physical
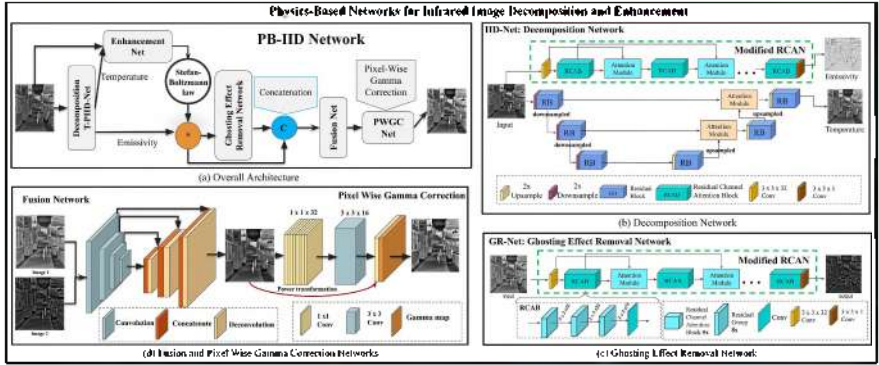
11

Figure 10: (a) Physics-driven deep PB-IID-Net Architecture for thermal image enhancement and analysis. It consists of several modules: (b) the decomposition IID-Net, enhancement net, (c) ghosting effect removal network (GR-Net), (d) fusion net, and a pixel-wise gamma correction (PWGC) network.
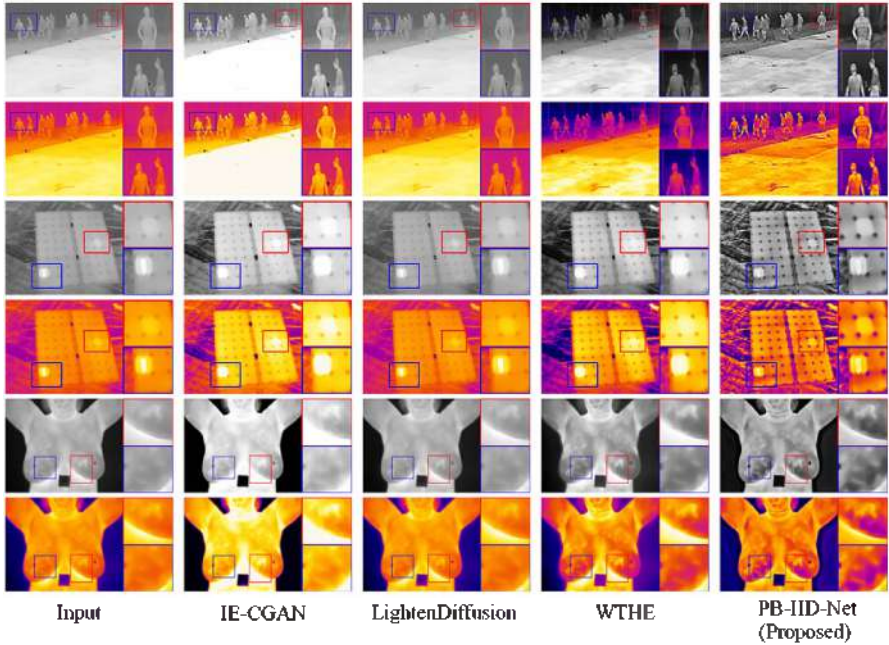

| Input | IE-CGAN | LightenDiffusion | WTHE | PB-IID-Net (Proposed) |

Figure 11: Generalizability of **PB-IID-Net**: preserving edges, textures, and contrast across diverse domains.

principles of thermal radiation and scene structure. Figure 10 illustrates the architecture of *PB-IID-Net*. The key contribution lies in using the extended Stefan-Boltzmann law to accurately decompose thermal images into temperature and emissivity components:

$$I = a \cdot \varepsilon \cdot T^K + b \qquad (4)$$

where I represents the final image intensity, ε denotes emissivity (the efficiency of a surface in emitting thermal energy), T is the absolute temperature, and a, b, and K are constants. This adaptation facilitates precise modeling of real-world surface characteristics in thermal imagery. The framework consists of sequential enhancement modules: **1.** Decomposition, which separates temperature and emissivity components using a newly proposed Laplacian pyramid network for improved detail visibility. **2.** Artifact suppression (GR-Net) eliminates ghost artifacts and sensor noise while restoring critical edges. **3.** Fusion Module merges enhanced images using pixel-wise gamma correction for optimal clarity. **4.** The final enhanced image is converted into visually interpretable RGB images via a novel pseudo-coloring mechanism with explicitly defined analytic color mapping (infrared and light-green) for enhanced human interpretation and analysis. Figure 11 highlights PB-IID-Net's capability to recover fine details and demonstrates improved interpretability through advanced colorization. Comprehensive evaluations across multiple datasets (LTIR, CVC-14, Autonomous Vehicles, Solar Panel, and Breast) confirm that PB-IID-Net significantly outperforms traditional and learning-based baselines across various thermal-specific metrics, including the recently proposed Local and Global Thermal Assessment (LGTA) [5]. The method consistently achieves superior object detection accuracy across different infrared spectra (Near, Mid, and Far), as shown in Table 3, underscoring its practical relevance in many real-world applications such as autonomous driving, medical imaging, and solar panel fault detection.
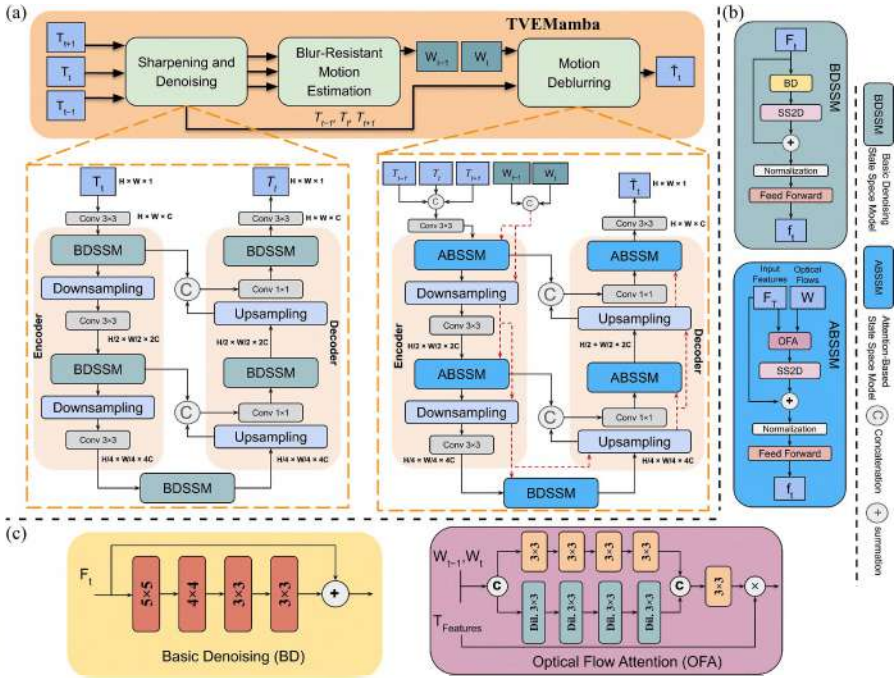

Figure 12: (a) Overall architecture of TVEMamba, (b) Basic denoising state space model and attention-based state space model, and (c) Basic denoising module and optical flow attention module.

Table 3: Detection performance (mAP) on various inputs and models

Measure	Orig. NIR	Orig. MIR	Orig. FIR	WTHE on NIR	WTHE on MIR	WTHE on FIR	PB IID Net on NIR $(k = 2)$	PB IID Net on MIR $(k = 2)$	PB IID Net on FIR $(k = 2)$	PB IID Net on NIR $(k = 4)$	PB IID Net on MIR $(k = 4)$	PB IID Net on FIR $(k = 4)$
mAP_{50} ↑	51.9	56.6	48.2	53.1	58.4	49.9	**60.5**	**63.8**	55.9	59.7	63.2	**56.4**
mAP_{75} ↑	14.8	20.1	16.3	15.3	22.1	18.5	**19.9**	**27.5**	20.7	18.4	27.1	**21.3**
$mAP_{50:95}$ ↑	22.6	25.1	22.5	23.5	26.8	24.1	**28.4**	**31.5**	27.7	26.9	30.9	**29.1**
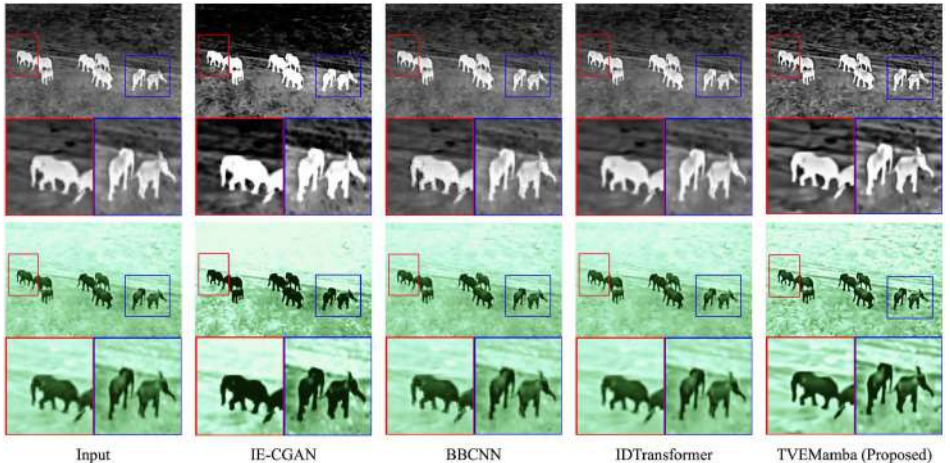
Figure 13: Qualitative comparison of motion artifact suppression on the wildlife BIRDSAI dataset.

Chapter 5 focuses on designing a Mamba-based framework for thermal video enhancement, specifically targeting core challenges such as low contrast, motion blur, sensor noise, and frame-to-frame inconsistencies that significantly hinder object tracking and detection performance. Existing SOTA methods often enhance contrast but at the cost of introducing temporal flicker, motion misalignment, and residual noise, limiting their practical reliability. To address these limitations, this chapter introduces **TVEMamba** [6], which effectively resolves these issues in thermal video processing. The overall pipeline of TVEMamba is illustrated in Figure 12. First, a sharpening and denoising network enhances sharpness and removes noise using a Mamba-based encoder-decoder structure integrated with the proposed Basic Denoising State Space Model (BDSSM), which includes a Basic Denoising (BD) block specifically designed for sharpening and denoising tasks. This structure efficiently captures both local details and global context. Next, a blur-resistant motion estimation module computes optical flows (W_{t-1} and W_t) between consecutive frames (T_{t-1}, T_t, T_{t+1}) using a global-to-local strategy for robustness in challenging conditions. Finally, the motion deblurring network integrates the newly proposed Attention-Based State Space Model (ABSSM) and Optical Flow Attention (OFA) modules into a Mamba-based encoder-decoder architecture, leveraging optical flow estimates to generate blur-free, corrected frames. These attention modules significantly enhance feature extraction and effective handling of motion-induced blur. Extensive evaluations across five datasets (BIRDSAI, FLIR, CAMEL, Autonomous Vehicles, and Solar Panels) demonstrate that TVEMamba surpasses

existing SOTA methods in multiple non-reference quality metrics, including EME, BDIM, DMTE, MDIMTE, LGTA, and BIE [5]. Figure 13 provides a detailed view of how TVEMamba preserves and refines complex image features. The improvements in edge sharpness, textural fidelity, and contrast balance are clearly visible, demonstrating the method's ability to recover essential scene details. Additionally, as shown in Table 4, TVEMamba enhances object detection accuracy by more than 15% compared to the original, severely degraded wildlife monitoring footage. These results underscore TVEMamba's practical efficacy in critical applications such as wildlife monitoring, autonomous driving, and UAV-based military operations, where precise and reliable object detection is essential.

Table 4: Object detection performance on the BIRDSAI dataset. $YOLO_1$ and Hyper-$YOLO_1$ models are trained on original datasets, and $YOLO_2$ and Hyper-$YOLO_2$ models are trained on enhanced datasets produced by the TVEMamba framework.

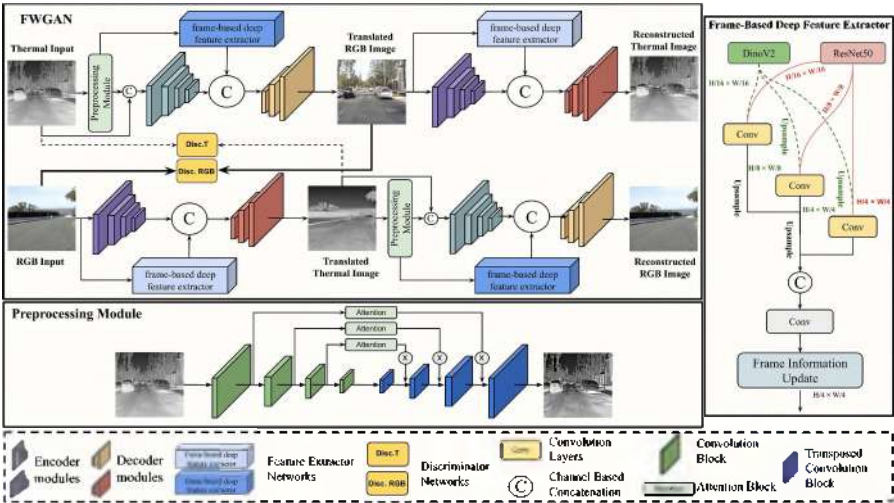| Classes	2		3		2	
Architecture	$YOLO_1$	$YOLO_2$	$YOLO_1$	$YOLO_2$	Hyper-$YOLO_1$	Hyper-$YOLO_2$
$mAP_{0.5}$	38.1	**44.2**	25.0	**29.7**	38.0	**43.9**
$mAP_{0.5:0.9}$	13.2	**16.8**	9.3	**10.9**	12.9	**16.4**

Figure 14: Overall Architecture of FWGAN for Thermal Image Translation and Analysis.

Chapter 6 presents a novel framework, **FWGAN**, designed for thermal image colorization, effectively translating thermal imagery into visually realistic RGB representations. FWGAN addresses current limitations, such as semantic distortions, temporal inconsistency, poor small-object handling, and insufficient texture preservation, particularly for objects where temperature variations can appear similar. The proposed framework extends our prior research [7] on thermal image colorization, which incorporates enhanced generative adversarial techniques with semantic-aware multitask learning for superior domain translation performance (Figure 14). Key innovations of the FWGAN include a
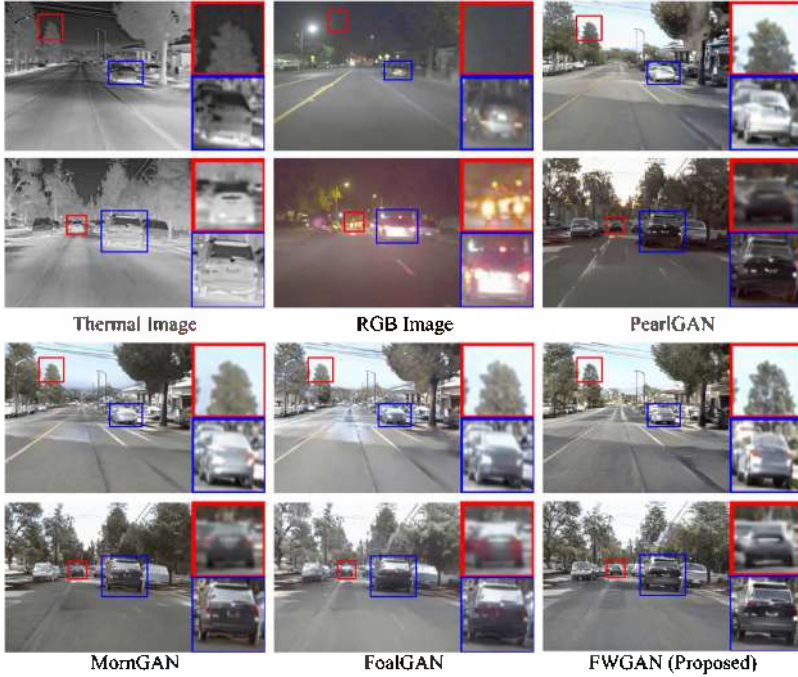
Figure 15: Qualitative comparison of FWGAN against SOTA methods for night-time TIR-to-RGB colorization.

Preprocessing Module that improves sharpness and reduces noise through an encoder-decoder structure, a Frame-based Deep Feature Extractor Module combining complementary semantic features from pretrained DINOv2 and ResNet50 backbones, and a Frame Information Update mechanism for maintaining temporal consistency across video frames. Furthermore, we propose an extended Weber Contrast Law-driven edge detection mechanism, significantly enhancing edge preservation:

$$f(\mathcal{L}) = 300 \cdot \left(\frac{L_{max} - L_{min}}{L_{min} + c} \right)^{0.01} \log \left(\frac{L_{max} - L_{min}}{L_{min} + c} \right) \tag{5}$$

where L_{max} and L_{min} denote maximum and minimum intensities in a local 3×3 region of thermal image L, and $c = 1$. Evaluations on FLIR and KAIST datasets demonstrate FWGAN's superior performance across key metrics (NIQE, BRISQUE, PIQE). Figure 15 highlights FWGAN's capability in realistic colorization and improved visibility of small objects and structural details. Compared to SOTA methods, FWGAN significantly enhances object detection accuracy by approximately 3.5%, underscoring its applicability in autonomous driving systems under challenging visibility conditions.

Chapter 7 concludes the thesis by summarizing the main contributions, discussing current limitations, and proposing potential future research on robust multi-modal perception in challenging environmental conditions.

Contributions

1. Developing **EOD-Net**, a specialized RGB image dehazing architecture that enhances visibility and object detection in hazy conditions, introducing a dual-branch dehazing system featuring the novel GLWF module and a specialized ABGC module for precise color restoration.

2. Developing **MTIE-Net**, a Mamba-based thermal image dehazing network introducing a novel convolution-based attention Enhancement and Denoising (ED) module in SSM for optimized image enhancement and a new hazy thermal image generation pipeline.

3. Developing **PB-IID-Net**, a physics-based thermal enhancement network that introduces extended Stefan-Boltzmann law for image decomposition/reconstruction, GR-Net for artifact reduction, and pseudo-coloring methods for better visual interpretation.

4. Developing **TVEMamba**, a Mamba-based thermal video enhancement framework, introducing specialized denoising and optical-flow estimation modules that address motion-induced blur and temporal inconsistencies, significantly improving detection accuracy.

5. Developing **FWGAN**, a Weber-law-driven GAN for TIR-to-RGB image translation, introducing a novel FDFEM Module and a preprocessing step, enhancing realism and semantic accuracy in translated images, crucial for improved RGB object detection performance in adverse conditions.

6. Evaluating real-world scenarios across diverse datasets demonstrates the effectiveness of these proposed methods when compared to existing SOTA techniques.

Practical Contributions of the Thesis

The practical contributions of this thesis are validated through extensive evaluations, highlighting the effectiveness of the proposed methods. The contributions include:

1. **Enhanced object detection for autonomous driving**: Evaluated on widely-used datasets including FLIR and KAIST, crucial for developing robust thermal object-detection systems. Significant improvements were observed in detection accuracy under fog and low-light conditions, essential for safe autonomous navigation.

2. **Improved wildlife monitoring capabilities**: Validated using the BIRDSAI dataset, commonly employed for evaluating UAV-based wildlife-detection systems. Enhanced detection accuracy facilitates more reliable conservation monitoring and anti-poaching efforts in challenging environments.

3. **Robust dehazing for challenging atmospheric conditions**: Validated against the top three NTIRE-2021 non-homogeneous dehazing methods, achieving superior quantitative results and improving visibility and detection accuracy in foggy conditions for environmental-monitoring applications.

4. **Reproducible research**: All generated datasets are publicly available[1], ensuring transparency and enabling further advances in robust object-detection methodology.

---

[1] https://github.com/SargisHovhannisyan/Object-Detection-in-Adverse-Weather

Potential Application Domains

The methodologies presented in this thesis have broad applicability across numerous safety-critical and high-value domains:

1. **Search and Rescue Operations**: Enhance visibility and consistency across video frames, crucial for detecting survivors and obstacles during UAV-assisted disaster response missions.

2. **Medical Imaging and Diagnostics**: Improve contrast and reduce distortions in thermal medical imagery, aiding early detection of diseases such as breast thermography and skin diagnostics.

3. **Military-based Systems**: Enhance visual clarity and object detection capabilities in adverse conditions, supporting critical military surveillance and tactical operations.

4. **Security Surveillance and Smart-City Monitoring**: Improve visual clarity in RGB and thermal images under low-light conditions, significantly boosting detection reliability and ensuring public safety.

List of Author's Publications

[1] S. Hovhannisyan, H. Gasparyan, S. Agaian, and A. Ghazaryan, "AED-Net: A single image dehazing," *IEEE access*, vol. 10, pp. 12 465–12 474, 2022, (Q1).

[2] S. Hovhannisyan, H. Gasparyan, and S. Agaian, "EOD-Net: Enhancing object detection in challenging weather conditions using an innovative end-to-end dehazing network," in *2023 Twelfth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, IEEE, 2023, pp. 1–6.

[3] S. Hovhannisyan, "Mamba-based thermal image dehazing," *Mathematical Problems of Computer Science*, vol. 62, pp. 126–144, 2024.

[4] S. Hovhannisyan, T. Trongtirakul, and S. Agaian, "Unified thermal image enhancement for object detection and visual clarity," *International Journal of Advances in Engineering and Management*, vol. 7, pp. 471–482, 2025.

[5] S. Hovhannisyan, S. Agaian, H. Ayunts, and T. Trongtirakul, "A new method for judging thermal image quality with applications," *Signal Processing*, vol. 229, p. 109 769, 2025, (Q1).

[6] S. Hovhannisyan, S. Agaian, K. Panetta, and A. Grigoryan, "Thermal video enhancement mamba: A novel approach to thermal video enhancement for real-world applications," *Information*, vol. 16, no. 2, p. 125, 2025, (Q2).

[7] S. Hovhannisyan and S. Agaian, "Thermal image enhancement and semantic multi task learning for accurate infrared colorization," *International Journal of Advances in Engineering and Management*, vol. 7, pp. 237–245, 2025.

Ամփոփում

Սարգիս Անդրեասի Հովհաննիսյան

Օբյեկտների հայտնաբերում անբարենպաստ եղանակային պայմաններում՝ օգտագործելով խորը ուսուցում և ջերմային-տեսանելի պատկերներ

Աշխատանքը **նվիրված** է օբյեկտների հայտնաբերման համակարգերի բարելավմանը և կայունության բարձրացմանը՝ հատկապես անբարենպաստ եղանակային պայմաններում (մառախուղ, ձյուն, մշուշ, ցածր լուսավորություն): Այս նպատակով մշակվել են խորը ուսուցման նոր մեթոդներ, որոնք օգտագործում են ինչպես գունավոր (RGB), այնպես էլ ջերմային (TIR) պատկերներ և տեսանյութեր: Կատարված գնահատումները հիմնավորում են մշակված լուծումների արդյունավետությունը և ցույց են տալիս դրանց կիրառելիությունը իրական կիրառական խնդիրներում:

Աշխատանքի **հիմնական նպատակներն** են՝

1. Կառուցել խորը ուսուցման նոր ցանցեր՝ ինչպես գունավոր (RGB), այնպես էլ ջերմային (TIR) պատկերներից մառախուղը կամ մշուշը հեռացնելու համար:

2. Ստեղծել խորը ուսուցման ցանցեր՝ ջերմային պատկերների և տեսանյութերի որակը բարելավելու համար (բարձրացնել հակադրությունը, նվազեցնել աղմուկը), որպեսզի հայտնաբերումը դառնա ավելի ճշգրիտ և կայուն:

3. Ստեղծել խորը ուսուցման նոր ցանց՝ ջերմային պատկերների գունավորման համար: Այն թույլ կտա օգտագործել օբյեկտների հայտնաբերման արդեն իսկ լայնորեն կիրառվող մեթոդները, խուսափելով ջերմային պատկերներով վերապատուցումից ինչը պահանջում է ջերմային տվյալների մեծ և պիտակավորված հավաքածուներ, որոնք կարող են սահմանափակ լինել:

4. Մանրամասն գնահատել առաջարկվող լուծումների արդյունավետությունը (հատկապես հայտնաբերման ճշգրտությունը)՝ դրանք փորձարկելով տարատեսակ իրական պայմաններում:

Առաջին **գլուխը** ներկայացնում է հետազոտության արդիականությունը, անբարենպաստ եղանակային օբյեկտների հայտնաբերման խնդիրը, աշխատանքի նպատակներն ու խնդիրները և ատենախոսության հիմնական նոր մոտեցումները:

Երկրորդ **գլուխը** ներկայացնում է խիտ մշուշոտ պայմաններում օբյեկտների հայտնաբերման արդյունավետությունը բարձրացնելու լուծումը՝ խորը ուսուցման նոր EOD-Net ցանցի միջոցով: Արդյունքները ցույց են տվել հայտնաբերման ճշգրտության զգալի բարելավում՝ **3-4%**-ից հասնելով մոտ **40%**-ի:

Երրորդ գլխում առաջարկվում է **MTIE-Net** անունով նոր ցանց (հիմնված «**Մամբա**» մոդելների վրա)՝ ջերմային պատկերներից մշուշը հեռացնելու համար: Այս մեթոդը

19

բարելավում է հայտնաբերման ճշգրտությունը **35%-ով**՝ համեմատած չմշակված մառախուղապատ ջերմային պատկերների հետ:

Չորրորդ գլխում ներկայացվում է ջերմային պատկերների որակը բարելավելու մեթոդ (մասսամբ հիմնված նրանց ֆիզիկական ներկայացման վրա)՝ **PB-IID-Net** ցանցի միջոցով: Այն կիրառելի է ինֆրակարմիր սպեկտրի ցանկացած տիրույթի համար և բարձրացնում է հայտնաբերման արդյունավետությունը ավելի քան **9%-ով**:

Հինգերորդ գլխում առաջարկվում է նոր **TVEMamba** ցանց (հիմնված **«Մամբա»** մոդելների վրա)՝ ջերմային տեսանյութերի որակը բարելավելու համար: Այն բարձրացնում է օբյեկտների հայտնաբերման ճշգրտությունը մոտ **15%-ով**:

Վեցերորդ գլխում մշակվել է **FWGAN** նոր ցանցը, որը ջերմային պատկերները վերածում է արհեստական գունավոր **(RGB)** պատկերների: Այս մեթոդը բարձրացնում է օբյեկտների հայտնաբերման ճշգրտությունը ավելի քան **3%-ով**:

Յոթերորդ գլխում ամփոփվում են ատենախոսության հիմնական արդյունքները և քննարկվում են անբարենպաստ եղանակային պայմաններում գործող հայտնաբերման համակարգերի հետագա բարելավման հնարավոր ուղղությունները:

Հիմնական կիրառական արդյունքներ

- Աշխատանքում մշակված մեթոդների արդյունավետությունը ստուգվել է համապարփակ փորձարկումներով, որոնք ցույց են տվել օբյեկտների հայտնաբերման համակարգերի ճշգրտության և կայունության զգալի բարելավում, հատկապես անբարենպաստ եղանակային պայմաններում:

- Մշակված մեթոդների գնահատումը կատարվել է ինքնավար տրանսպորտային համակարգերի ստանդարտացված ջերմային **(TIR)** պատկերների հավաքածուներով **(FLIR, KAIST)**, որոնք լայնորեն օգտագործվում են իրական համակարգերի նախագծման և գնահատման համար: Արդյունքները ցույց են տվել հայտնաբերման զգալի աճ մշուշի և ցածր լուսավորության պայմաններում:

- Առաջարկված մեթոդների արդյունավետությունը ստուգվել է **BIRDSAI** տեսանյութերի հավաքածուի միջոցով, որը նախատեսված է անօդաչու թռչող սարքերով վայրի կենդանիների (փղերի) դիտարկման համար: Այն լայնորեն կիրառվում է հակաորսագողության և բնապահպանական տեսահսկման իրական համակարգերի գնահատման համար:

- Աշխատանքում օգտագործված բոլոր գեներացված և կիրառված տվյալների հավաքածուները հրապարակվել են համացանցում՝ ապահովելով ստացված արդյունքների ամբողջական վերարտադրելիությունը, ինչն աջակցում է հետագա հետազոտություններին և մեթոդների գործնական կիրառությանը:

Заключение

Оганнисян Саргис Андреасович

Обнаружение объектов в неблагоприятных погодных условиях с использованием методов глубокого обучения, а также видимых и тепловых изображений

Работа посвящена улучшению и повышению устойчивости систем обнаружения объектов, особенно в неблагоприятных погодных условиях (туман, дым, дымка, условия низкой освещенности). С этой целью были разработаны новые методы глубокого обучения, использующие как цветные (**RGB**), так и тепловые (**TIR**) изображения и видеозаписи. Проведенные оценки обосновывают эффективность разработанных решений и демонстрируют их применимость в реальных прикладных задачах.

Основные цели работы:

1. Построить новые сети глубокого обучения для удаления тумана или дымки как с цветных (**RGB**), так и с тепловых (**TIR**) изображений.

2. Создать сети глубокого обучения для улучшения качества тепловых изображений и видео (повысить контрастность, снизить уровень шума), чтобы обнаружение стало более точным и устойчивым.

3. Создать новую сеть глубокого обучения для колоризации тепловых изображений. Это позволит использовать уже широко применяемые методы обнаружения объектов, избегая переобучения на тепловых изображениях, что требует больших аннотированных наборов тепловых данных, которые могут быть ограничены.

4. Детально оценить эффективность предлагаемых решений (особенно точность обнаружения), тестируя их в разнообразных реальных условиях.

Первая глава представляет актуальность исследования, проблему обнаружения объектов в неблагоприятных погодных условиях, цели и задачи работы, а также основные новые подходы диссертации.

Вторая глава представляет решение для повышения эффективности обнаружения объектов в условиях густой дымки с помощью новой сети глубокого обучения EOD-Net. Результаты показали значительное улучшение **точности** обнаружения: с 3-4% до примерно 40%.

В третьей главе предлагается новая сеть под названием **MTIE-Net** (основанная на моделях «Мамба») для удаления дымки с тепловых изображений. Этот метод улучшает точность обнаружения на **35%** по сравнению с необработанными тепловыми изображениями в условиях тумана/дымки.

В четвертой главе представляется метод улучшения качества тепловых изображений (частично основанный на их физических характеристиках) с помощью сети **PB-IID-Net**. Он применим для любого диапазона инфракрасного спектра и повышает эффективность обнаружения более чем на **9%**.

В пятой главе предлагается новая сеть **TVEMamba** (основанная на моделях «Мамба») для улучшения качества тепловых видеозаписей. Она повышает точность обнаружения объектов примерно на **15%**.

В шестой главе разработана новая сеть **FWGAN**, которая преобразует тепловые изображения в искусственно окрашенные (**RGB**) изображения. Этот метод повышает точность обнаружения объектов более чем на **3%**.

В седьмой главе обобщаются основные результаты диссертации и обсуждаются возможные направления дальнейших исследований, связанные с развитием систем обнаружения, работающих в неблагоприятных погодных условиях.

Основные практические результаты

- Эффективность разработанных в работе методов была проверена путем всесторонних испытаний, которые показали значительное улучшение точности и устойчивости систем обнаружения объектов, особенно в неблагоприятных погодных условиях.

- Оценка разработанных методов проводилась на стандартизированных наборах тепловых (**TIR**) изображений для автономных транспортных систем (**FLIR**, **KAIST**), которые широко используются для проектирования и оценки реальных систем. Результаты показали значительный рост показателей обнаружения в условиях дымки и низкой освещенности.

- Эффективность предложенных методов была проверена с помощью набора видеоданных **BIRDSAI**, предназначенного для наблюдения за дикими животными (слонами) с использованием беспилотных летательных аппаратов. Он широко применяется для оценки реальных систем борьбы с браконьерством и экологического мониторинга.

- Все сгенерированные и использованные в работе наборы данных были опубликованы в интернете, обеспечивая полную воспроизводимость полученных результатов, что способствует дальнейшим исследованиям и практическому применению методов.